A peer-reviewed open-access journal

NeoBiota

Advancing research on alien species and biological invasions

NeoBiota 104: 339–359 (2025)
DOI: 10.3897/neobiota.104.162310

Research Article

From detection to action—a proposed workflow to ensure first reports of alien species from molecular analyses are acted upon

Laura Fernández Winzer[1,2,3], Katelyn T. Faulkner[1,4], Trudy Paap[5], John R. U. Wilson[1,2]

1 *South African National Biodiversity Institute, Kirstenbosch Research Centre, Cape Town, South Africa*
2 *Department of Botany and Zoology, Centre for Invasion Biology, Stellenbosch University, Stellenbosch, South Africa*
3 *School of Natural Sciences, Macquarie University, Sydney, Australia*
4 *Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa*
5 *Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa*
Corresponding author: Laura Fernández Winzer (laura.f.winzer@gmail.com)

OPEN ACCESS

This article is part of:
Developing lists of alien taxa in the Global South: workflows, protocols, processes, and experiences
Edited by John Wilson, Michele Dechoum, Katelyn Faulkner, Barbara Langdon, Shyama Pagad, Aníbal Pauchard, Hanno Seebens, Tsungai Zengeya, Silvia Ziller

Academic editor: Hanno Seebens
Received: 18 June 2025
Accepted: 7 November 2025
Published: 17 December 2025

Citation: Fernández Winzer L, Faulkner KT, Paap T, Wilson JRU (2025) From detection to action—a proposed workflow to ensure first reports of alien species from molecular analyses are acted upon. NeoBiota 104: 339–359. https://doi.org/10.3897/neobiota.104.162310

Copyright: © Laura Fernández Winzer et al. This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Abstract

Gaps between detection and action often hinder timely biosecurity responses. For example, the polyphagous shot hole borer (*Euwallacea fornicatus*) was first recorded in South Africa in 2012 through DNA barcoding, but action only followed a field observation in 2017. Here we present a simple, generalisable workflow that aims to ensure molecular observations are evaluated and flagged for further action. Using South Africa as a case study, 10,084 records from the 'Barcode of Life Data System' (BOLD) for South Africa were compared with three datasets: a watch list of species of concern (400 species); a list of high-risk agricultural pests detected at the border but not yet established (218 species); and the Botanical Database of Southern Africa (BODATSA, 68,153 records), which includes alien plant species found outside of cultivation. We identified four species on the watch list or pest list that appear to be present in South Africa based on molecular observations, flagged four species as possible additions to the list of alien plants outside cultivation, and identified several discrepancies between lists that should be investigated and resolved. This, we believe, demonstrates the potential for the workflow to guide prioritisation of surveillance and biosecurity actions. The value of the workflow, however, depends on the curation of the underlying lists, how regularly it is applied, and institutional uptake. The workflow should thus be seen as a potential tool to improve the flow of information from detection to action within the context of an integrated and responsive biosecurity system.

Key words: Biosecurity, BOLD, DNA barcoding, EDRR, prohibited species, species identification, species of concern, taxonomic standardisation

Introduction

Detecting an alien species soon after it has been introduced is likely to mean more options are available for management (Reaser et al. 2020). At early stages, populations are typically small and geographically limited, making containment and eradication more feasible [for example, early detection of red imported fire ant (*Solenopsis invicta*) in Australia enabled successful eradication efforts in several incursion sites (Wylie et al. 2020)]. As a result, large-scale harmful invasions can potentially be avoided (IPBES 2023). Two aspects of biosecurity frameworks are particularly important to achieve this: first, surveillance systems need to be improved

so alien species introductions are detected; and second, a system is needed to flag detections that should be acted upon (Liebhold et al. 2016).

An infamous example from South Africa is the polyphagous shot hole borer (PSHB), an ambrosia beetle that together with its fungal symbiont causes dieback and mortality of a diverse range of woody plants. This beetle was detected through a sentinel project at Kwa-Zulu Natal Botanical Gardens in 2017 (Paap et al. 2018), however, it soon became clear that it was already widely established across much of the country (Paap et al. 2020). The 2017 detection raised national awareness, and by 2021 the PSHB was known to be present in eight out of South Africa's nine provinces, affecting many native, ornamental, and commercially important tree species (van Rooyen et al. 2021; de Wit et al. 2022; Townsend et al. 2025). Retrospective analysis revealed that the PSHB had been present in South Africa as early as 2012, when it was sampled as part of a barcoding project (van Rooyen et al. 2021). Whether effective action could have been taken in 2012 is not known, but it is concerning that there were five years between the first detection and the subsequent detection that was acted upon.

Globally, DNA barcoding has emerged as a powerful tool for the detection of alien taxa, providing molecular evidence of their presence in a given region while supporting biosecurity efforts (Armstrong and Ball 2005; Hamelin 2012; Hanner et al. 2009; Madden et al. 2019). The method involves sequencing a standardised region of an organism's genome and comparing it against reference databases. The resulting matches can identify species with high confidence (see Baena-Bejarano et al. 2023). DNA barcoding has proven particularly useful for identifying cryptic species or when morphological identification is challenging due to specimen damage or when specimens are only available for life stages that are not ideal for identification (as they are missing important diagnostic characters) (Saccaggi and Ueckermann 2024). Despite the potential for reducing the time between detection and response, the integration of molecular barcode data into national and regional biosecurity frameworks remains limited (Armstrong and Ball 2005).

The Barcode of Life Data System (BOLD; https://boldsystems.org/) is an open-access DNA barcode database that provides a centralised platform for barcode data (Ratnasingham et al. 2024). Each record is assigned a Barcode Index Number (BIN), a system for clustering DNA barcode sequences into operational taxonomic units that typically correspond to species (Ratnasingham and Hebert 2013). As of May 2025, BOLD contained 16.5 million public records from around the world, representing 1.3 million species across various kingdoms (e.g., Animalia, Bacteria, Chromista, Fungi, and Plantae). BOLD's global coverage allows researchers to use this tool to detect and identify species across a broad range of geographic regions and taxonomic groups, supporting biodiversity monitoring and conservation efforts, and increasingly contributing to biosecurity applications worldwide (Hernández-Triana et al. 2019; Comia and Morris 2024).

Importantly, integrative approaches combining molecular and morphological data have been shown to enhance identification accuracy, highlighting the limitations of relying solely on morphology or genetics (Yang et al. 2022). Many barcode-based detections, particularly those derived from environmental DNA (eDNA) or uncurated sequence submissions, require follow-up verification to confirm species identity and presence (Burian et al. 2021). Misidentifications, contamination, or incomplete reference databases can lead to false positives or uncertain records (Goldberg et al. 2016). It is therefore essential that molecular

detections are assessed within a structured verification framework that integrates taxonomic expertise, curated reference data, and, where possible, field confirmation. Many countries still rely heavily on morphological identification, which is often labour-intensive and limited by the availability of taxonomic expertise (Darling and Blum 2007). This gap in molecular-based surveillance systems leaves many regions vulnerable to undetected introductions of alien taxa, particularly in biodiversity hotspots like South Africa, where rapid and accurate identification is crucial (Sethusa et al. 2014; van Wilgen et al. 2020).

This study aims to address this gap by developing a simple, generalisable workflow that cross-references barcode data from BOLD with lists of taxa that have been flagged as of concern. Specifically we used data of BOLD specimens collected in South Africa and compared those with three datasets: (1) a national watch list of alien species of concern; (2) a list of high-risk agricultural pests intercepted at borders but not yet detected within the country; and (3) the Botanical Database of Southern Africa (BODATSA), which includes alien plants outside of cultivation. By identifying taxa that have been detected during molecular studies but that are not yet recognised or managed as alien species in South Africa (i.e., included in the aforementioned lists), we aim to flag potential biosecurity risks to inform action and improve existing species lists for future surveillance. This workflow represents a step towards integrating molecular data into practical biosecurity applications, with the potential to be adapted for use in other regions facing similar challenges.

Methods

To compare the species lists and identify potential undocumented alien species, we developed a workflow comprising three main steps: (1) acquisition, initial inspection, and preparation of the input datasets; (2) taxonomic standardisation of species names across datasets [following Faulkner (in press)]; and (3) pairwise comparison of standardised lists (Fig. 1, Suppl. material 1: fig. S1). The workflow is implemented in R (v4.4.2; R Core Team 2024) and relies on several R packages. These include rgbif (Chamberlain et al. 2025), rWCVP (Brown et al. 2023), and rWCVPdata (Govaerts et al. 2021; Govaerts 2024); which provide access to the taxonomic backbones of the Global Biodiversity Information Facility (GBIF) and the World Checklist of Vascular Plants (WCVP), according to which species names are standardised (step 2 of the workflow).

Species lists

The primary input dataset was derived from the Barcode of Life Data System (BOLD v4. https://boldsystems.org/), providing molecular evidence of species presence in South Africa. This list was compared to three reference datasets to flag species of concern. Below, we describe each input dataset and their limitations.

Barcode of Life Database (BOLD)

The BOLD database (Ratnasingham and Hebert 2007; Ratnasingham et al. 2024) is a global, open-access DNA barcoding platform that includes animals, plants, fungi, bacteria, and protists. Available metadata may include collection date and location, collector identity, specimen storage information, and associated media.

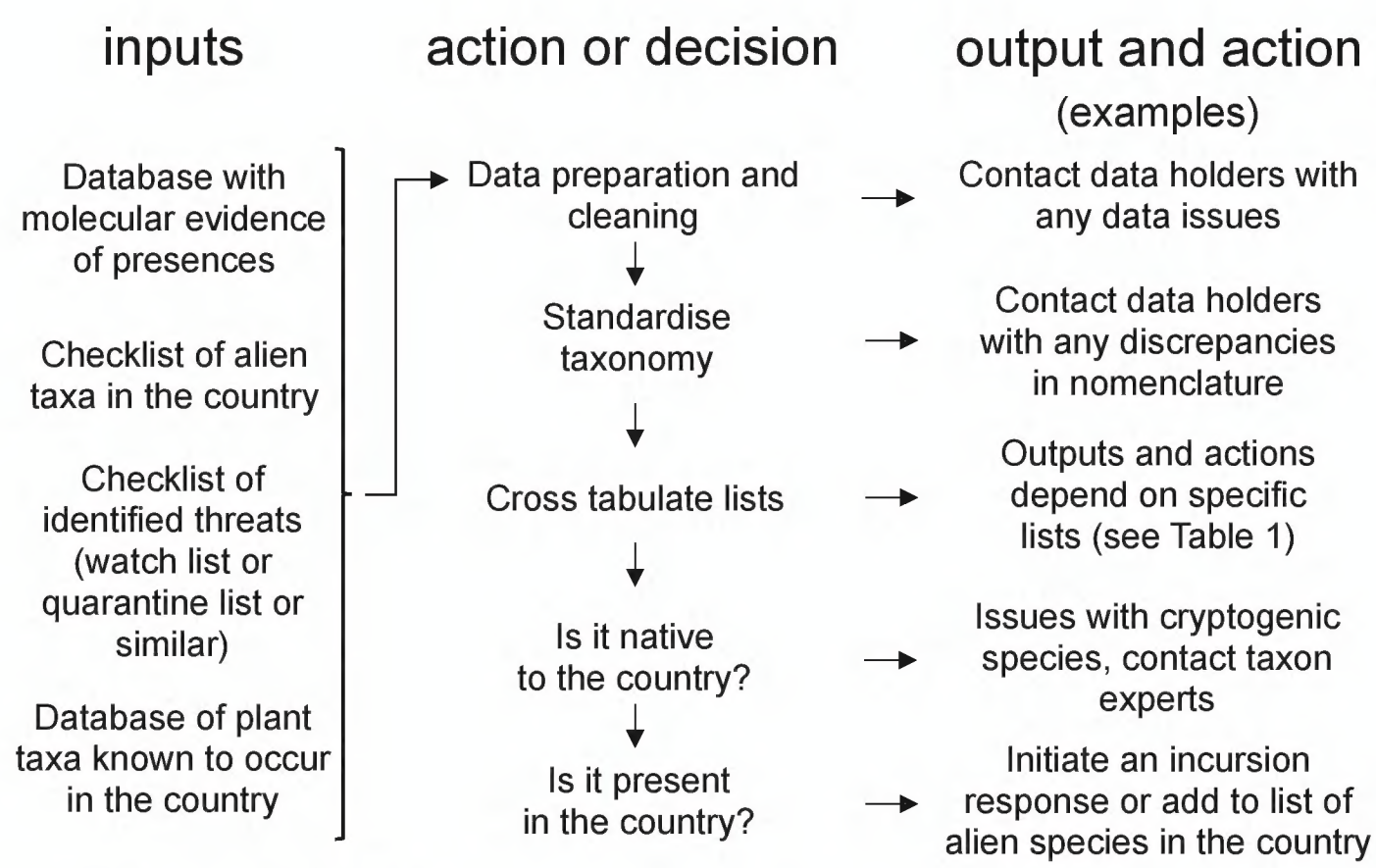
inputs
action or decision
output and action
(examples)
Database with molecular evidence of presences
Checklist of alien taxa in the country
Checklist of identified threats (watch list or quarantine list or similar)
Database of plant taxa known to occur in the country
Data preparation and cleaning
Standardise taxonomy
Cross tabulate lists
Is it native to the country?
Is it present in the country?
Contact data holders with any data issues
Contact data holders with any discrepancies in nomenclature
Outputs and actions depend on specific lists (see Table 1)
Issues with cryptogenic species, contact taxon experts
Initiate an incursion response or add to list of alien species in the country

Figure 1. An overview of the workflow for comparing country specific species lists with molecular databases.

Critically, the database does not indicate whether a taxon is native to the region where the specimen was collected, posing challenges for detecting alien species (discussed below). Records lacking DNA sequence data were excluded from this workflow. Data were downloaded on 22 November 2023, with a South African subset obtained by applying the search filter "South Africa"[geo] within the BOLD data portal to extract all records associated with the country.

Watch list

This list, compiled by Faulkner et al. (2014), includes 400 alien species considered to pose a high risk of invasion but not known to be present in South Africa as of 2013. It comprises animals, plants (including algae and ferns), and other taxa such as fungi, oomycetes and other micro-organisms. As the list is static, some species on it have been detected since the analysis was completed in early 2013 (e.g., *Austropuccinia psidii*; Roux et al. 2013).

If a species on the watch list is found to have a DNA barcode collected from samples in South Africa, a response is warranted. The evidence supporting the BOLD record (including species identification, collection location, voucher specimens, and the number of independent records) should be assessed. Where possible, additional verification such as contacting the authors, re-sampling, or examining images and specimens from herbaria or museums should be undertaken. Checks should be made to see if the species was already known in the country (and the watch list updated accordingly), and/or field confirmation should potentially trigger an incursion response.

High risk pest list ('inspections list')

The list, from Saccaggi et al. (2021; available on figshare: Terblanche et al. 2021), contains 218 alien taxa intercepted on plant imports between 1994 and 2019 but not known to occur in South Africa. Taxa on this list includes mainly invertebrates (Phylum Arthropoda and Nematoda), but also viruses, fungi, and bacteria.

If a species on this list is found to have a DNA barcode collected from samples in South Africa, then the response should be the same as described for the watch list. Like the watch list, this list is not updated, so verification would be needed to confirm a species has not been reported since 2021.

Plants of Southern Africa (BODATSA)

BODATSA (also known as NewPOSA; http://posa.sanbi.org/) is a living database curated by the South African National Biodiversity Institute (SANBI), documenting the presence and status [e.g., native (referred to as 'indigenous'), naturalised, invasive] of plant species in South Africa. BODATSA includes vascular plants and many non-vascular plant groups such as Bryophyta (mosses) and Marchantiophyta (liverworts) but it does not include algae. It does not include alien taxa only recorded within cultivation (i.e., horticultural species may be included in BOLD but not in BODATSA). The version used here was downloaded on 6th May 2025 and contained 68,153 rows (SANBI 2025). The most recent version should be downloaded when performing this analysis, with the date of download recorded (as specified in R script 2, Suppl. material 1).

If an alien plant is present on BOLD with a DNA barcode but is not listed in BODATSA, the first step is to assess the evidence of the BOLD record by checking whether it includes collection details such as coordinates or a locality name. If the species is only present in South Africa under cultivation, it will not appear in BODATSA (though monitoring for signs of escape from cultivation and naturalisation may still be advisable separate to this workflow). However, if the BOLD record corresponds to individuals outside of cultivation, then the taxon should be flagged for potential inclusion in BODATSA and an evaluation of the need for an incursion response made (Wilson et al. 2013). Where appropriate, physical specimens should be collected and curated as evidence of the alien taxon's presence in the country.

Workflow

The workflow is reproducible, adjustable, transparent, and based on existing templates (Reyserhove et al. 2020; Seebens et al. 2020), with R scripts freely available (see Suppl. material 1, https://github.com/KatelynFaulkner/bold-first-reports-workflow).

Data preparation (R script 1)

The living database (BOLD) and static datasets (watch list and pest list) were cleaned and stored for reuse. Details are provided below. Note that BODATSA is not subjected to taxonomic standardisation because it serves as the taxonomic backbone for plants in South Africa. As such, it is not included in the first step of data preparation. Instead, it is processed during the second step (and corresponding script), where it is prepared both for use as the taxonomic reference and for list comparisons in step 3.

BOLD

To ensure there is molecular evidence of a species' presence, we retained only BOLD records with DNA sequences. This was assessed using the 'nucleotides' column in the "Combined" file export (which includes both Specimen and Sequence data). Records with blank, dashed, or missing values in this column were excluded.

Note that presence of a BIN (Barcode Index Number) indicates sequencing, but some sequenced records lack BINs, so the 'nucleotides' field is more inclusive. Unnecessary columns and duplicated species names were removed, as well as records with no data for species name. The 'species_name' column was renamed to 'verbatimScientificName' following Darwin Core standards. The list was saved as a .csv file.

The version used initially contained 498,735 rows. After removing 81,761 blank entries, 334 dashed lines, 9 records from Israel (mislabelled under South Africa), and the duplicated species, 10,084 entries remained.

Watch list

From the original dataset (884 taxa), we selected the 400 species designated as "Watch list" in the 'Final.designation' column. The 'Species name' field was renamed to 'verbatimScientificName', unnecessary columns were removed, and the list was saved as a .csv file.

Inspection list

The dataset "Metadata of contaminants on SA plant imports 1994–2019.csv" (878 taxa; Terblanche et al. 2021) was downloaded from Figshare. We selected species recorded as "absent" in South Africa (column 'SA.occurenceStatus.current'), removed unnecessary columns, and renamed the 'species' column to 'verbatimScientificName'. The list was saved as a .csv file.

Taxonomic standardisation (R script 2)

To ensure consistent taxonomy, we followed the standardisation workflow developed by Faulkner (in press), which was developed for South African alien species lists (Zengeya et al. 2025b). The watch list, inspection list, and list from BOLD were standardised using this workflow (https://github.com/KatelynFaulkner/rsa-ans-workflow). As per Faulkner (in press), BODATSA was used for plant names and the GBIF backbone was used for the names of other taxa. If a plant taxon was not found in BODATSA, the taxonomic backbone of the World Checklist of Vascular Plants (via POWO) was used. The automated workflow was adapted so that the taxonomic standardisation of the lists was iterative. As described in Faulkner (in press) manual inspection of the outputs of the automated workflow was required to resolve ambiguous cases (e.g., homonyms) and ensure clean outputs (Murray et al. 2017, see details in Suppl. material 1). When uncertain or conflicting taxonomic information was encountered, species names were verified by consulting taxonomic experts, reviewing source publications, and contacting relevant database curators (e.g., BODATSA) to confirm name validity and species identity. Duplicates were removed to retain one row per species. The outputs were saved as .csv files, with the column 'scientificName' containing the taxonomically standardised taxon names.

As previously mentioned, the BODATSA list is not subjected to taxonomic standardisation because it serves as the taxonomic backbone for plants in South Africa. However, it still requires preparation to be used effectively in this role. We did not remove native or endemic species at this stage to avoid misclassifying native taxa in BOLD as aliens during list comparison. The formatting in some columns was

standardised (e.g., double spaces removed). BODATSA lists the scientific name across multiple columns ('Genus', 'Sp1', 'Rank1', 'Sp2'), which were merged to a single 'scientificName' field which contained the scientific name with authorship and date information. Duplicated names in 'scientificName' were removed as well as unnecessary columns. Taxa labelled as "misapplied" in the 'TaxStat' column, indicating taxonomic misidentifications or records not considered valid for South Africa, were also excluded from the dataset. The list was saved as a .csv file.

List comparison (R script 3)

Lists (i.e., scientificName columns) were compared in pairs, with BOLD always serving as the reference. While the comparison between BOLD and the watch list, and BOLD and the inspection list flags species present in both lists, the comparison of BOLD and BODATSA flags species present only in BOLD (i.e., missing from BODATSA).

BOLD vs Watch & Inspection lists

The complete BOLD list (10,084 records) was compared to the 400 watch list taxa and 185 pest species. Any matches were flagged as potential new detections. Because the watch and pest lists are static, all matches required literature checks to ensure the species had not already been reported in South Africa after 2013 (watch list) or 2021 (pest list).

BOLD vs BODATSA

For this comparison, only plant taxa from BOLD were used. The aim was to identify species in BOLD not recorded in BODATSA. Such mismatches may represent unrecognised or undocumented plant introductions and/or naturalisations. Subsequent checks were required to determine if these are indeed new records for South Africa.

Results

Number of species per species list

The BOLD dataset downloaded on 22 November 2023 contained 498,735 records for South Africa, of which 11.5% were identified to species level (56,357 records), representing 10,084 species. After taxonomic standardisation, 801 synonyms (verbatim name was a synonym of an accepted name that is included in the scientific name column), and 313 duplicates for scientific name were identified.

The watch list from Faulkner et al. (2014) originally included 400 species. The taxonomic standardisation detected no duplicates, but encountered 44 synonyms. The high-risk pest list from Saccaggi et al. (2021) comprised 218 high-risk pest species, which was reduced to 185 following taxonomic checks and removal of duplicates, with five synonyms and four scientific name duplicates.

Finally, the BODATSA list, downloaded on 6th May 2025, originally contained 68,153 records, which were reduced to 67,105 species after data preparation and cleaning (i.e., 1,048 misapplied names were removed).

Taxonomic standardisations

The inspection list after going through the steps of data preparation and taxonomic standardisation (scripts 1 and 2) resulted in 37 records (out of 218) that needed to be manually checked before moving to the next step of list comparison. These were 'possible error flags' related to either uncertain (six), doubtful (one) or unmatched taxon names (30) [see Suppl. material 1 and Faulkner (in press) for more details].

The watch list after the taxonomic standardisation contained ten possible error flags to manually check out of 400 records (four uncertain, five doubtful and one unmatched).

Finally, the BOLD list with 10,084 records had 1,589 possible error flags to be manually checked (16% of the total records: 193 uncertain, 61 doubtful, and 1,366 unmatched). Depending on the input dataset, this step can be highly time-consuming, particularly when numerous flags require manual verification, however, the manual work is greatly reduced by implementing the automated workflow. Importantly, unmatched records were predominantly taxa identified only to the genus level (72%), which could not be resolved to species level during taxonomic standardisation. These records were retained at the genus level, as the entire genus could potentially be absent from the country—even if species-level identification was not available. Certain genera are known to include pests or invasive species, so flagging the presence of a new genus in the country was considered important. It should be noted that because these genus-level records do not require species-level verification, the proportion of records requiring full manual checking was substantially reduced from 16% to 4.3%. After all flags were manually verified, we proceeded with the list comparisons.

List comparisons

BOLD vs Watch List

Seven species were shared between BOLD (10,084 species) and the Faulkner et al. (2014) watch list (400 species; Table 1; Suppl. material 2). These were the fungal pathogen *Austropuccinia psidii*, four invertebrates (*Aedes aegypti*, *Cinara cupressi*, *Culex quinquefasciatus* and *Octolasion tyrtaeum*), the grey wolf *Canis lupus*, and a cultivated plant (*Terminalia catappa*).

Austropuccinia psidii is a highly invasive fungus affecting plant species within the Myrtaceae family. Known as the causal agent of myrtle rust, it was first detected in South Africa in 2013 (Roux et al. 2013), after Faulkner et al. (2014) had analysed the data for their study. This, therefore, does not represent a new introduction to the country (note that in the watch list, Roux et al. 2013 and in BOLD the species appears under the synonym *Puccinia psidii*). Likewise, *A. aegypti* and *C. cupressi* were found to be in South Africa after Faulkner et al. (2014) was published (see Guarido et al. 2021 and Wondafrash et al. 2024 respectively). *Canis lupus* is also not flagged because, although there are 54 records in BOLD for this species, they pertain to domestic dogs (*Canis lupus familiaris* or *Canis familiaris*) and not grey wolves, which are the species of concern on the watch list. Nonetheless this highlights the value of the workflow for updating watch lists and similar.

Table 1. Proposed management actions for when molecular evidence of presence is found for alien taxa listed under different check list types. The numbers of species are based on the South African examples. The priorities for action are with reference specifically to this workflow: 1) immediate action is required; 2) some action should be taken, but this is not urgent (e.g., could be part of annual updates); 3) continued actions may be advisable, but not specifically emerging from this workflow; and 4) no action needed.

Check list type	Molecular evidence of presence	Listed on check list	Priority for action	Proposed action(s)	Number of species in South African example	Example
Watch list (Faulkner et al. 2014)	TRUE	TRUE	1	Species are identified as threats if they were not previously recorded as present in the country. Verify BOLD record evidence, confirm presence in the country, update watch list, and initiate an incursion response.	7 initially but after manual check, 3 flagged	*Culex quinquefasciatus*, *Octolasion tyrtaeum*, and *Terminalia catappa*
	TRUE	FALSE	4	No action needed. Species are present in the country but are not identified as significant threats (many are likely native).	10,077	*Achroia grisella*
	FALSE	TRUE	3	Species are identified as threats, but there is no evidence of presence in the country. Continue efforts to prevent introductions.	393	*Cronartium ribicola*
	FALSE	FALSE	4	No action needed; species are recorded neither as present nor a threat.	NA	NA
Quarantine pest list (Saccaggi et al. 2021)	TRUE	TRUE	1	Species are identified as threats and were not previously recorded as present in the country. Verify BOLD record evidence, confirm presence in the country, update quarantine list, and initiate an emergency pest response plan.	2 initially but after manual check, 1 flagged	*Cartodere constricta*
	TRUE	FALSE	4	No action needed. Species are present in the country, but are not identified as quarantine pests (many are likely native).	10,082	*Culex annulioris*
	FALSE	TRUE	3	Species are identified as quarantine pests, but there is no evidence of presence in the country. Continue efforts to prevent future introductions and monitor for signs of presence post-border.	183	*Aceria tulipae*
	FALSE	FALSE	4	No action needed; species are recorded neither as present nor a threat.	NA	NA
National plant checklist (BODATSA)	TRUE	TRUE	4	No action needed; species are recorded as present in both databases.	3035	*Pinus halepensis*
	TRUE	FALSE	1	Verify BOLD record evidence, incorporate species into checklist, and, as necessary, collect and curate physical specimen as evidence of the alien taxon's presence in the country.	205 from which 131 were found to be alien, 9 flagged	*Ficus microcarpa*, *Jacobaea maritima*, and *Terminalia catappa*
	FALSE	TRUE	2	Arrange for a barcode (or similar) to be collected so there is molecular evidence of the alien taxon's presence in the country.	1,763	*Colocasia esculenta*
	FALSE	FALSE	4	No action needed; species are not recorded as present in both databases.	NA	NA

Therefore, three species were flagged for action through the BOLD–Watch list comparison: (1) *Culex quinquefasciatus* (the southern house mosquito), (2) *Octolasion tyrtaeum* (the Woodland white worm, an earthworm), and (3) *Terminalia catappa* (the tropical almond). The first two species require immediate attention, as they are recognised threats and had not been previously recorded as present in South Africa. For these taxa, the validity of the BOLD record must be verified, their presence in the country confirmed, the watch list updated accordingly, and an incursion response initiated (for more detail on these species see Suppl. material 2). For *T. catappa*, which current records indicate is present only in cultivation, ongoing monitoring is recommended to ensure it does not escape and establish. If monitoring confirms that the species has spread beyond cultivation and formed a naturalised population, it should then be flagged for inclusion in BODATSA and assessed for an incursion response. In such cases, physical specimens should be collected and curated as supporting evidence.

BOLD vs High-Risk Pest List

Two taxa were shared between the BOLD dataset (10,084 species) and the high-risk pest list (185 species) (Table 1; Suppl. material 3): the plaster beetle

Cartodere constricta and the fruit fly *Drosophila immigrans*. After checking the species manually, *Drosophila immigrans* was found to already be known to be present in South Africa (De Araujo et al. 2019). Therefore, only one species was flagged: *C. constricta*. According to GBIF (https://www.gbif.org/species/1044381, accessed 11 May 2025), this beetle has been introduced into six countries. However, no evidence of environmental or economic impact has been recorded to date.

BOLD vs BODATSA

Comparison of the plant species in BOLD (3,384 species) with the BODATSA list (67,105 species) revealed 323 species unique to BOLD (see Table 1; Suppl. material 4). Of these, 144 were algae, which are not included in BODATSA, leaving 179 taxa for further assessment: 161 vascular plants (Phylum Tracheophyta) and 18 non-vascular plants [10 mosses (Bryophyta) and eight liverworts (Marchantiophyta)]. Of the remaining taxa, 47 were determined to be native (according to POWO) but missing from BODATSA, while one was considered data deficient. This left 131 taxa confirmed as alien to South Africa. From this, nine species were flagged: *Combretum indicum*, *Conocarpus erectus*, *Ficus microcarpa*, *Jacobaea maritima*, *Olea europaea* subsp. *europaea*, *Phormium tenax*, *Terminalia catappa* (see above), *Voacanga africana*, and *Yucca gloriosa*.

While several flagged species are confirmed invaders elsewhere, their current status in South Africa varies (for details see Suppl. material 1). Some are presently under cultivation (*C. indicum*, *C. erectus*, *O. europaea* subsp. *europaea*, *T. catappa* and *V. africana*) others are recorded only from offshore islands (*F. microcarpa*, *P. tenax* and *Y. gloriosa*), and one lacks locality data (*J. maritima*). All may already be spreading and therefore warrant further scrutiny. For those four not under cultivation, once the validity of the BOLD records is confirmed, they should be incorporated into BODATSA, and physical specimens should be collected and curated as evidence of their presence in the country.

Discussion

The proposed workflow offers a practical tool for bridging the gap between molecular detection and management response. By integrating different types of species lists and standardising taxonomic information, it fosters more consistent communication between researchers, biosecurity agencies, and decision-makers (Gatto et al. 2013; Pagad et al. 2018; Sandall et al. 2023). It also enables more efficient triage of flagged taxa and encourages routine surveillance using molecular records (van Rees et al. 2022; Pocock et al. 2024).

Comparison between the species lists flagged four taxa that warrant further investigation. These include three species from the watch list and one from the list of high-risk quarantine pests. Given their prior designation as species of concern, risk analyses should be prioritised and an immediate response considered. Additional plant taxa were also flagged from the comparison with BODATSA. These records require field verification to determine whether they represent new introductions or are already established but undocumented; in either case these taxa should be evaluated for potential inclusion in BODATSA.

Importantly, not all flagged species necessarily represent novel incursions. Taxonomic inconsistencies, synonymy, or unresolved native versus alien status can lead to apparent mismatches across datasets (Pyšek et al. 2013; Essl et al. 2018). This highlights the importance of treating each flagged species as a candidate for expert scrutiny, rather than assuming it to be a new alien species. Misidentification or misclassification can generate false positives, and careful interpretation is necessary before initiating management actions (Cristescu and Hebert 2018; Hayes 2021). In this context, we will notify BODATSA of all species that appear as native to South Africa in POWO but are not listed in BODATSA, to support efforts towards consistency between datasets. Discrepancies between authoritative sources such as POWO and BODATSA present a significant hurdle for end users, complicating the interpretation of presence and nativity. Resolving such inconsistencies is essential to improve the reliability of species checklists and support evidence-based biosecurity responses.

The decision to remove misapplied names from BODATSA (>1,000 records) was deliberate to prevent confusion during the comparison process. Misapplied names represent instances where species were incorrectly identified in the past, leading to false assumptions about their presence in the country. By eliminating these records, the workflow ensures that any flagged species appearing in BOLD for South Africa can be confidently considered as potential incursions or established populations. For example, the name *Acacia retinoides* was historically misapplied to naturalised populations of *A. provincialis* in South Africa, a discrepancy that was only resolved recently (Magona et al. 2018). Another decision that needed to be made was how to handle taxa with ambiguous identifications—such as those labelled with aff. (affinis), cf. (confer), or nr (near). These were manually reviewed during the standardisation process. Such qualifiers typically indicate uncertainty at the species level, suggesting that the specimen is similar to, but not confidently assignable as, a particular species. To reflect this uncertainty and avoid incorrectly flagging an entire genus, these records were standardised to the format "Genus sp." This approach ensures that any potential flagging or follow-up actions remain appropriately focused on a single, unidentified species within the genus, rather than generalising to the entire taxonomic group.

While the present analysis focused on a relatively small number of flagged taxa, the workflow is designed to be scalable. In broader contexts such as continental scale studies or countries with more extensive molecular surveillance programs, the number of alien species detected could be substantially higher. Applying the workflow at such scales would require greater computational resources, and automated quality-control steps. However, one aspect that is unlikely to be completely resolved through automated processes is taxonomic standardisation, as not all the procedures required can be automated (Murray et al. 2017). In the analysis presented here 37 records (17%) on the inspection list, 10 records on the watch list (2.5%), and 1,589 records (16%) on the BOLD list required manual checking, although 1,150 of these records in BOLD were taxa identified to genus-level and thus did not require these manual checks. Implementing automated taxonomic standardisation did significantly reduce the time and effort involved in this step, but the burden of manual review would increase if the analysis was scaled up. One particular issue is that the species names in the three datasets that required taxonomic standardisation appear without authorship and date information. This greatly impacts the ability of automated tools

to resolve species names, as taxonomic homonyms (taxon names that are identical in spelling, but belong to different taxa) cannot be resolved through automated processes unless authorship and date information are provided (Grenié et al. 2022). Scaling up the analysis would, therefore, require coordinated taxonomic expertise; we recommend that datasets, such as BOLD, routinely provide species names with authorship and date information in future.

For taxa detected exclusively through eDNA, it is crucial to establish evidence of physical presence before assuming incursion. Burgess et al. (2021), in their study on *Phytophthora* in Australia, emphasised that while eDNA and metabarcoding are valuable for species detection, physical verification remains necessary. They recommend that eDNA detections be followed up with traditional isolation methods to confirm the presence of live organisms. In alignment with ISPM8 guidelines (IPPC Secretariat 2021), taxa detected solely by molecular methods should not be considered present without accompanying biological confirmation. This precaution guards against false positives that could mislead biosecurity efforts (Darling and Mahon 2011; Rees et al. 2014). Extending this principle beyond *Phytophthora* to other taxa is essential for ensuring robust detection and response strategies.

An important consideration when using BOLD as a detection tool is its reliance on complete metadata for optimal effectiveness. For BOLD records to be truly informative, it is crucial that users provide detailed data, including the collection location, date, and any available images (see Suppl. material 1: table S2, S3 for a complete list). These ancillary data fields significantly facilitate field verification and incursion response efforts. Pitcher et al. (2024) emphasise that BOLD records are particularly valuable when accompanied by comprehensive ancillary data and underscore the need for meticulous curation practices and for as many data fields to be filled in as possible.

Challenges remain, particularly regarding invertebrate detection. Invertebrates are often underrepresented in barcoding databases, not only in South Africa but globally (Pitcher et al. 2024). Boykin et al. (2012) discuss how the lack of comprehensive barcoding for invasive insects poses risks to biosecurity. Addressing these gaps requires ongoing projects like the QBOL project in Europe, which focuses on gathering DNA barcoding data of quarantine organisms (see Bonants et al. 2010), and multiple efforts from South Africa. For example, in 2011, the International Barcode of Life Project (iBOL) node was established in South Africa, followed by the launch of the Urban DNA Barcode Project in KwaZulu-Natal. Additionally, funding mechanisms like the National Research Foundation's Foundational Biodiversity Information Programme have supported these developments (da Silva and Willows-Munro 2016; Pitcher et al. 2024). These projects have made significant progress in expanding barcoding efforts for invasive and native species alike. However, continuous funding and support for DNA barcoding remain critical to sustain these initiatives and enhance detection capabilities.

Even when clear evidence of invasive species presence is available—such as in the case of the PSHB in South Africa—mobilising a rapid and coordinated response can be difficult. Decision-makers and the public often require visible or large-scale impacts before taking action, by which point management options may be limited (Simberloff et al. 2013). A similar delay was observed in Australia during the early spread of myrtle rust (*Austropuccinia psidii*), despite warnings from scientists (Carnegie and Pegg 2018). The tendency to delay action until impacts are widely visible illustrates the socio-political and economic barriers to effective incursion response (Hester and Bland 2024).

Recommendations

The workflow presented here has the potential to assist with efforts to reduce the rates of introduction and establishment of known or potentially invasive alien species [paraphrased from Target 6 of the Kunming-Montreal Global Biodiversity Framework (CBD 2022)]. The workflow is not a substitute for preventative measures, but it complements them by enhancing incursion response (Wilson et al. 2017; Clarke et al. 2023). Once a species has entered the country, rapid identification and assessment are critical to reduce the risk of spread, and workflows like the one we propose can support those efforts. For the workflow to be effective in practice, however, it would need to be institutionalised and included as part of standard operating procedures. Surveillance operations need to be in place, the data gathered through such exercises need to be picked up and applied to this workflow, and if risks are flagged there needs to be a clear reporting chain to someone who can initiate an incursion response. In this sense we believe the workflow is a tool to assist with integrated governance (Roy et al. 2024).

We recommend applying this workflow at least annually, given that the BOLD database is regularly updated with new records from diverse molecular studies (Ratnasingham et al. 2024). Although many of these studies are not focused specifically on alien species, they often generate data with incidental biosecurity value (Comtet et al. 2015). Regular application of the workflow ensures that newly deposited records are not overlooked, and it strengthens the early warning capacity of national biosecurity systems.

In order to institutionalise the workflow in South Africa, its implementation could build on existing structures and mandates. Species lists and their updates could be curated and housed in GBIF, where lists of alien species present in South Africa and the Prince Edward Islands are already published or in preparation (Fernández Winzer et al. 2024; Zengeya et al. 2025b). SANBI currently coordinates national reporting on biological invasions through the National Status Report on Biological Invasions (produced every three years) and maintains national alien species lists (Zengeya et al. 2025a). The long-term goal is to update these lists annually, and automating elements of the workflow could help achieve this. The Department of Forestry, Fisheries and the Environment (DFFE) could use the system to receive alerts of potential new records, enabling rapid assessment and communication to relevant taxonomic experts within SANBI and to the incursion response team (Wilson et al. 2013). Strengthening articulation with South Africa's Border Management Authority would also be critical to ensure that detections from molecular data are integrated with interception records and surveillance activities at points of entry.

For the workflow to be implemented in other countries or regions, key baseline information resources are needed. Some countries may not yet have comprehensive alien species lists; in such cases, the first step would be to compile national or regional inventories of alien taxa and watch lists of species of concern. Notably, an automated tool to develop watch lists, based on the procedure set out in Faulkner et al. (2014), has recently been developed (Ivey et al. 2025), and should assist with this. Once such lists are available, the workflow can be used to compare molecular records (e.g., from BOLD) with those lists to detect potential new incursions. Each country would also need to identify institutions responsible for hosting and maintaining these lists, and to designate the authorities or teams responsible for verifying flagged records and

initiating an incursion response. The workflow thus can serve as an important tool within any country or biosecurity system to flag records of concern and to identify issues of misalignment between different data sources.

Acknowledgements

Francois Roets and Davina Saccaggi for useful discussion on the first steps of the paper when looking for lists of high-risk pests, Mark Robertson for constructive comments..

Additional information

Conflict of interest

KF and JW were part of the editorial team for the special issue 'Developing lists of alien taxa in the Global South: workflows, protocols, processes, and experiences', but they took no part in the peer review and decision-making process for this paper.

Ethical statement

No ethical statement was reported.

Use of AI

No use of AI was reported.

Funding

This work was supported by the Centre for Invasion Biology (CIB) at Stellenbosch University. The South African Department of Forestry, Fisheries and the Environment (DFFE) are thanked for funding, noting that this publication does not necessarily represent the views or opinions of DFFE or its employees. KF and JW also received funding from the B-cubed project (Biodiversity Building Blocks for policy) which is funded by the European Union's Horizon Europe Research and Innovation Programme (ID No 101059592). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them. TP was funded by a University of Pretoria Postdoctoral Research Fellowship.

Author contributions

T.P. and J.R.U.W. conceived the idea; L.F.W. obtained the species lists, performed preliminary analysis and led the writing of the manuscript; K.T.F. developed the R scripts. All coauthors provided insightful advice, valuable feedback on the manuscript and agreed to submit the final version of the manuscript.

Author ORCIDs

Laura Fernández Winzer https://orcid.org/0000-0003-0675-9117
Katelyn T. Faulkner https://orcid.org/0000-0002-3955-353X
Trudy Paap https://orcid.org/0000-0003-1364-4350
John R. U. Wilson https://orcid.org/0000-0003-0174-3239

Data availability

All raw data files supporting the findings of this study are available on Zenodo (https://doi.org/10.5281/zenodo.17501334). The R scripts underpinning the analysis reported in this paper are provided in the Supplementary Information and are also accessible via a GitHub repository (https://github.com/KatelynFaulkner/bold-first-reports-workflow).

References

Armstrong KF, Ball SL (2005) DNA barcodes for biosecurity: Invasive species identification. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 360: 1813–1823. https://doi.org/10.1098/rstb.2005.1713

Baena-Bejarano N, Reina C, Martínez Revelo DE, Medina CA, Tovar E, Uribe-Soto S, et al. (2023) Taxonomic identification accuracy from BOLD and GenBank databases using over a thousand insect DNA barcodes from Colombia. PLoS ONE 18(4): e0277379. https://doi.org/10.1371/journal.pone.0277379

Bonants P, Groenewald E, Rasplus JY, Maes M, De Vos P, Frey J, Boonham N, Nicolaisen M, Bertacini A, Robert V, Barker I (2010) QBOL: A new EU project focusing on DNA barcoding of quarantine organisms. Bulletin OEPP. EPPO Bulletin. European and Mediterranean Plant Protection Organisation 40(1): 30–33. https://doi.org/10.1111/j.1365-2338.2009.02350.x

Boykin LM, Armstrong KF, Kubatko L, De Barro P (2012) Species delimitation and global biosecurity. Evolutionary Bioinformatics Online 8: 1–37. https://doi.org/10.4137/EBO.S8532

Brown MJM, Walker BE, Black N, Govaerts RHA, Ondo I, Turner R, Nic Lughadha E (2023) rWCVP: A companion R package for the World Checklist of Vascular Plants. The New Phytologist 240(4): 1355–1365. https://doi.org/10.1111/nph.18919

Burgess TI, Edwards J, Drenth A, Massenbauer T, Cunnington J, Mostowfizadeh-Ghalamfarsa R, Dinh Q, Liew EC, White D, Scott P, Barber PA (2021) Current status of *Phytophthora* in Australia. Persoonia - Molecular Phylogeny and Evolution of Fungi. 47: 151–177. https://doi.org/10.3767/persoonia.2021.47.05

Burian A, Mauvisseau Q, Bulling M, Domisch S, Qian S, Sweet M (2021) Improving the reliability of eDNA data interpretation. Molecular Ecology Resources 21(5): 1422–1433. https://doi.org/10.1111/1755-0998.13367

Carnegie AJ, Pegg GS (2018) Lessons from the incursion of myrtle rust in Australia. Annual Review of Phytopathology 56: 457–478. https://doi.org/10.1146/annurev-phyto-080516-035256

CBD (2022) Convention on Biological Biodiversity, Kunming-Montreal Global Biodiversity Framework. https://www.cbd.int/doc/decisions/cop-15/cop-15-dec-04-en.pdf

Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L, Ram K (2025) rgbif: interface to the Global Biodiversity Information Facility API, R package version 3.8.0. https://cran.r-project.org/package=rgbif

Clarke LJ, Shaw JD, Suter L, Atalah J, Bergstrom DM, Biersma E, Convey P, Greve M, Holland O, Houghton MJ, Hughes KA, Johnston EL, King CK, McCarthy AH, McGaughran A, Pertierra LR, Robinson SA, Sherman CDH, Stark JS, Stevens MI, Strugnell JM, von Ammon U, Wilson NG, Zaiko A, MacDonald AJ (2023) An expert-driven framework for applying eDNA tools to improve biosecurity in the Antarctic. Management of Biological Invasions : International Journal of Applied Research on Biological Invasions 14(3): 379–402. https://doi.org/10.3391/mbi.2023.14.3.01

Comia CF, Morris MR (2024) Beyond BOLD: Using DNA barcoding to identify Prussian carp (*Carassius gibelio* Bloch, 1782) in southern Alberta, Canada. BioInvasions Records 13(2): 497–513. https://doi.org/10.3391/bir.2024.13.2.15

Comtet T, Sandionigi A, Viard F, Casiraghi M (2015) DNA (meta) barcoding of biological invasions: A powerful tool to elucidate invasion processes and help managing aliens. Biological Invasions 17: 905–922. https://doi.org/10.1007/s10530-015-0854-y

Cristescu ME, Hebert PD (2018) Uses and misuses of environmental DNA in biodiversity science and conservation. Annual Review of Ecology, Evolution, and Systematics 49(1): 209–230. https://doi.org/10.1146/annurev-ecolsys-110617-062306

da Silva JM, Willows-Munro S (2016) A review of over a decade of DNA barcoding in South Africa: A faunal perspective. African Zoology 51: 1–12. https://doi.org/10.1080/15627020.2016.1151377

Darling JA, Blum MJ (2007) DNA-based methods for monitoring invasive species: A review and prospectus. Biological Invasions 9: 751–765. https://doi.org/10.1007/s10530-006-9079-4

Darling JA, Mahon AR (2011) From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. Environmental Research 111: 978–988. https://doi.org/10.1016/j.envres.2011.02.001

De Araujo LI, Karsten M, Terblanche JS (2019) Three new Drosophilidae species records for South Africa. Bothalia-African Biodiversity & Conservation 49(1): 1-5. https://doi.org/10.4102/abc.v49i1.2429

de Wit MP, Crookes DJ, Blignaut JN, de Beer ZW, Paap T, Roets F, van der Merwe C, van Wilgen BW, Richardson DM (2022) An assessment of the potential economic impacts of the invasive Polyphagous Shot Hole Borer (Coleoptera: Curculionidae) in South Africa. Journal of Economic Entomology 115: 1076–1086. https://doi.org/10.1093/jee/toac061

Essl F, Bacher S, Genovesi P, Hulme PE, Jeschke JM, Katsanevakis S, et al. (2018) Which taxa are alien? Criteria, applications, and uncertainties. Bioscience 68: 496–509. https://doi.org/10.1093/biosci/biy057

Faulkner KT (in press) An automated workflow to standardise taxon names for South African alien species lists. African Biodiversity & Conservation.

Faulkner KT, Robertson MP, Rouget M, Wilson JR (2014) A simple, rapid methodology for developing invasive species watch lists. Biological Conservation 179: 25–32. https://doi.org/10.1016/j.biocon.2014.08.014

Fernández Winzer L, Greve M, le Roux P, Faulkner K, Wilson J, Pagad S (2024) Protected Areas - Global Register of Introduced and Invasive Species - Prince Edward Island and Marion Island, South Africa. Version 1.5. Invasive Species Specialist Group ISSG. Checklist dataset. https://cloud.gbif.org/griis/resource?r=pa-prince-edward-islands&v=1.5

Gatto F, Katsanevakis S, Vandekerkhove J, Zenetos A, Cardoso AC (2013) Evaluation of online information sources on alien species in Europe: The need of harmonization and integration. Environmental Management 51: 1137–1146. https://doi.org/10.1007/s00267-013-0042-8

Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Spear SF, McKee A, Oyler-McCance SJ, Cornman RS, Laramie MB (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in Ecology and Evolution 7(11): 1299–1307. https://doi.org/10.1111/2041-210X.12595

Govaerts R (2024) WCVP: World Checklist of Vascular Plants. Facilitated by the Royal Botanic Gardens, Kew, version 13.

Govaerts R, Nic Lughadha E, Black N, Turner R, Paton A (2021) The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. Scientific Data 8(1): 215. https://doi.org/10.1038/s41597-021-00997-6

Grenié M, Berti E, Carvajal-Quintero J, Dädlow GML, Sagouis A, Winter M (2022) Harmonizing taxon names in biodiversity data: A review of tools, databases and best practices. Methods in Ecology and Evolution 2022: 1–14. https://doi.org/10.1111/2041-210X.13802

Guarido MM, Riddin MA, Johnson T, Braack LEO, Schrama M, Gorsich EE, Brooke BD, Almeida APG, Venter M (2021) *Aedes* species (Diptera: Culicidae) ecological and host feeding patterns in the north-eastern parts of South Africa, 2014–2018. Parasites & Vectors 14: 339. https://doi.org/10.1186/s13071-021-04845-9

Hamelin RC (2012) Contributions of genomics to forest pathology. Canadian Journal of Plant Pathology 34: 20–28. https://doi.org/10.1080/07060661.2012.665389

Hanner RH, Lima J, Floyd R (2009) DNA barcoding and its relevance to pests, plants and biological control. ISHS Acta Horticulturae 823: 41–48. https://doi.org/10.17660/ActaHortic.2009.823.3

Hayes KA (2021) Taxonomic shortcuts lead to long delays in species discovery, delineation, and identification. Biological Invasions 23: 1285–1292. https://doi.org/10.1007/s10530-020-02438-8

Hernández-Triana LM, Brugman VA, Nikolova NI, Ruiz-Arrondo I, Barrero E, Thorne L, de Marco MF, Krüger A, Lumley S, Johnson N, Fooks AR (2019) DNA barcoding of British mosquitoes (Diptera, Culicidae) to support species identification, discovery of cryptic genetic diversity and monitoring invasive species. ZooKeys 832: 57–76. https://doi.org/10.3897/zookeys.832.32257

Hester SM, Bland LM (2024) Prepare, Respond, and Recover: Selecting Immediate and Long-Term Strategies to Manage Invasions. Biosecurity. Taylor & Francis, 18 pp. https://doi.org/10.1201/9781003253204-10

IPBES (2023) Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES secretariat, Bonn, Germany. https://doi.org/10.5281/zenodo.7430682

IPPC Secretariat (2021) Determination of pest status in an area. International Standard for Phytosanitary Measures No. 8. Rome. FAO on behalf of the Secretariat of the International Plant Protection Convention. https://openknowledge.fao.org/server/api/core/bitstreams/2a296eef-4cf4-447a-b84a-4d8d03a3bafc/content

Ivey PJ, Faulkner KT, Miller J, van Steenderen CJM (2025) *WatchListR*: a tool for developing watchlists of invasive species to inform biosecurity decision-making. NeoBiota 104: 163–177. https://doi.org/10.3897/neobiota.104.163164

Liebhold AM, Berec L, Brockerhoff EG, Epanchin-Niell RS, Hastings A, Herms DA, Kean JM, McCullough DG, Suckling DM, Tobin PC, Yamanaka T (2016) Eradication of invading insect populations: From concepts to applications. Annual Review of Entomology 61(1): 335–352. https://doi.org/10.1146/annurev-ento-010715-023809

Madden MJL, Young RG, Brown JW, Miller SE, Frewin AJ, Hanner RH (2019) Using DNA barcoding to improve invasive pest identification at U.S. ports-of-entry. PLoS ONE 14: e0222291. https://doi.org/10.1371/journal.pone.0222291

Magona N, Richardson DM, Le Roux JJ, Kritzinger-Klopper S, Wilson JR (2018) Even well-studied groups of alien species might be poorly inventoried: Australian *Acacia* species in South Africa as a case study. NeoBiota 39: 1–29. https://doi.org/10.3897/neobiota.39.23135

Murray BR, Martin LJ, Phillips ML, Pyšek P (2017) Taxonomic perils and pitfalls of dataset assembly in ecology: A case study of the naturalized Asteraceae in Australia. NeoBiota 34: 1–20. https://doi.org/10.3897/neobiota.34.11139

Paap T, De Beer ZW, Migliorini D, Nel WJ, Wingfield MJ (2018) The polyphagous shot hole borer (PSHB) and its fungal symbiont *Fusarium euwallaceae*: A new invasion in South Africa. Australasian Plant Pathology 47: 231–237. https://doi.org/10.1007/s13313-018-0545-0

Paap T, Wingfield MJ, De Beer ZW, Roets F (2020) Lessons from a major pest invasion: The polyphagous shot hole borer in South Africa. South African Journal of Science 116: 11/12. https://doi.org/10.17159/sajs.2020/8757

Pagad S, Genovesi P, Carnevali L, Schigel D, McGeoch MA (2018) Introducing the global register of introduced and invasive species. Scientific Data 23: 1–2. https://doi.org/10.1038/sdata.2017.202

Pitcher T, Ndaba A, Jacobs A, Hamer M, Janion-Scheepers C (2024) DNA barcoding of alien invertebrates and biological control agents in South Africa: A review. African Entomology 32: e19290. https://doi.org/10.17159/2254-8854/2024/a19290

Pocock MJ, Adriaens T, Bertolino S, Eschen R, Essl F, Hulme PE, Jeschke JM, Roy HE, Teixeira H, De Groot M (2024) Citizen science is a vital partnership for invasive alien species management and research. IScience 19: 27(1). https://doi.org/10.1016/j.isci.2023.108623

Pyšek P, Hulme PE, Meyerson LA, Smith GF, Boatwright JS, Crouch NR, Figueiredo E, Foxcroft LC, Jarošík V, Richardson DM, Suda J, Wilson JRU (2013) Hitting the right target: Taxonomic challenges for, and of, plant invasions. AoB Plants 5(0): plt042. https://doi.org/10.1093/aobpla/plt042

R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7: 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Ratnasingham S, Hebert PDN (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. PLoS ONE 8(8): e66213. https://doi.org/10.1371/journal.pone.0066213

Ratnasingham S, Wei C, Chan D, Agda J, Agda J, Ballesteros-Mejia L, Ait Boutou H, El Bastami ZM, Ma E, Manjunath R, Rea D, Ho C, Telfer A, McKeowan J, Rahulan M, Steinke C, Dorsheimer J, Milton M, Hebert PDN (2024) BOLD v4: A Centralized Bioinformatics Platform for DNA-Based Biodiversity Data. DNA Barcoding: Methods and Protocols. Chapter 26. Springer US, New York, 403-441. https://doi.org/10.1007/978-1-0716-3581-0_26

Reaser JK, Burgiel SW, Kirkey J, Brantley KA, Veatch SD, Burgos-Rodríguez J (2020) The early detection of and rapid response (EDRR) to invasive species: A conceptual framework and federal capacities assessment. Biological Invasions 22: 1–19. https://doi.org/10.1007/s10530-019-02156-w

Rees HC, Maddison BC, Middleditch DJ, Patmore JR, Gough KC (2014) The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. Journal of Applied Ecology 51: 1450–1459. https://doi.org/10.1111/1365-2664.12306

Reyserhove L, Desmet P, Oldoni D, Adriaens T, Strubbe D, Davis AJ, et al. (2020) A checklist recipe: Making species data open and FAIR. Database : The Journal of Biological Databases and Curation. https://doi.org/10.1093/database/baaa084

Roux J, Greyling I, Coutinho TA, Verleur M, Wingfeld MJ (2013) The Myrtle rust pathogen, *Puccinia psidii*, discovered in Africa. IMA Fungus 4: 155–159. https://doi.org/10.5598/imafungus.2013.04.01.14

Roy HE, Pauchard A, Stoett PJ, Renard Truong T, Meyerson LA, Bacher S, Galil BS, Hulme PE, Ikeda T, Kavileveettil S, McGeoch MA, Nuñez MA, Ordonez A, Rahlao SJ, Schwindt E, Seebens H, Sheppard AW, Vandvik V, Aleksanyan A, Ansong M, August T, Blanchard R, Brugnoli E, Bukombe JK, Bwalya B, Byun C, Camacho-Cervantes M, Cassey P, Castillo ML, Courchamp F, Dehnen-Schmutz K, Zenni RD, Egawa C, Essl F, Fayvush G, Fernandez RD, Fernandez M, Foxcroft LC, Genovesi P, Groom QJ, González AI, Helm A, Herrera I, Hiremath AJ, Howard PL, Hui C, Ikegami M, Keskin E, Koyama A, Ksenofontov S, Lenzner B, Lipinskaya T, Lockwood JL, Mangwa DC, Martinou AF, McDermott SM, Morales CL, Müllerová J, Mungi NA, Munishi LK, Ojaveer H, Pagad SN, Pallewatta NPKTS, Peacock LR, Per E, Pergl J, Preda C, Pyšek P, Rai RK, Ricciardi A, Richardson DM, Riley S, Rono BJ, Ryan-Colton E, Saeedi H, Shrestha BB, Simberloff D, Tawake A, Tricarico E, Vanderhoeven S, Vicente J, Vilà M, Wanzala W, Werenkraut V, Weyl OLF, Wilson JRU, Xavier RO, Ziller SR (2024) Curbing the major and growing threats from invasive alien species is urgent and achievable. Nature Ecology & Evolution 8: 1216–1223. https://doi.org/10.1038/s41559-024-02412-w

Saccaggi DL, Ueckermann EA (2024) The problem of taxonomic uncertainty in biosecurity: South African mite interceptions as an example. Acaralogia 64(2): 363–369. https://doi.org/10.24349/top1-r59v

Saccaggi DL, Arendse M, Wilson JR, Terblanche JS (2021) Contaminant organisms recorded on plant product imports to South Africa 1994–2019. Scientific Data 8(1): 83. https://doi.org/10.1038/s41597-021-00869-z

SANBI (2025) South African National Biodiversity Institute, Botanical Database of Southern Africa (BODATSA) [https://posa.sanbi.org/sanbi/Explore] [downloaded 06 May 2025]

Sandall EL, Maureaud AA, Guralnick R, McGeoch MA, Sica YV, Rogan MS, Booher DB, Edwards R, Franz N, Ingenloff K, Lucas M (2023) A globally integrated structure of taxonomy to support biodiversity science and conservation. Trends in Ecology & Evolution 38(12): 1143–1153. https://doi.org/10.1016/j.tree.2023.08.004

Seebens H, Clarke DA, Groom Q, Wilson JR, García-Berthou E, Kühn I, et al. (2020) A workflow for standardising and integrating alien species distribution data. NeoBiota 59: 39–59. https://doi.org/10.3897/neobiota.59.53578

Sethusa MT, Yessoufou K, Van der Bank M, Van der Bank H, Millar IM, Jacobs A (2014) DNA barcode efficacy for the identification of economically important scale insects (Hemiptera: Coccoidea) in South Africa. African Entomology 22(2): 257–266. https://doi.org/10.4001/003.022.0218

Simberloff D, Martin JL, Genovesi P, Maris V, Wardle DA, Aronson J, Courchamp F, Galil B, García-Berthou E, Pascal M, Pyšek P (2013) Impacts of biological invasions: What's what and the way forward. Trends in Ecology & Evolution 28(1): 58–66. https://doi.org/10.1016/j.tree.2012.07.013

Terblanche J, Saccaggi DL, Arendse M, Wilson JRU (2021) Contaminant organisms recorded on plant product imports to South Africa 1994-2019. figshare. Collection 8: 83. https://doi.org/10.1038/s41597-021-00869-z

Townsend G, Hill M, Hurley BP, Roets F (2025) Escalating threat: Increasing impact of the polyphagous shot hole borer beetle, *Euwallacea fornicatus*, in nearly all major South African forest types. Biological Invasions 27: 88. https://doi.org/10.1007/s10530-025-03551-2

van Rees CB, Hand BK, Carter SC, Bargeron C, Cline TJ, Daniel W, Ferrante JA, Gaddis K, Hunter ME, Jarnevich CS, McGeoch MA (2022) A framework to integrate innovations in invasion science for proactive management. Biological Reviews of the Cambridge Philosophical Society 97(4): 1712–1735. https://doi.org/10.1111/brv.12859

van Rooyen E, Paap T, De Beer W, Townsend G, Fell S, Nel WJ, Morgan S, Hill M, Gonzalez A, Roets F (2021) The polyphagous shot hole borer beetle: current status of a perfect invader in South Africa. South African Journal of Science 10: 117(11-12). https://doi.org/10.17159/sajs.2021/9736

van Wilgen BW, Measey J, Richardson DM, Wilson JR, Zengeya TA (2020) Biological Invasions in South Africa. Springer Nature, 975 pp. https://doi.org/10.1007/978-3-030-32394-3

Wilson JRU, Ivey P, Manyama P, Nänni I (2013) A new national unit for invasive species detection, assessment and eradication planning. South African Journal of Science 109: 0111. [0113 pp.] https://doi.org/10.1590/sajs.2013/20120111

Wilson JR, Panetta FD, Lindgren C (2017) Detecting and responding to alien plant incursions. Cambridge University Press, 265 pp. https://doi.org/10.1017/CBO9781316155318

Wondafrash M, Wingfield MJ, Hurley BP, Slippers B, Mutitu EK, Jenya H, Paap T (2024) DNA sequence data confirms the presence of two closely related cypress-feeding aphid species on African cypress (*Widdringtonia* spp.) in South Africa. Southern Forests 86(4): 278–285. https://doi.org/10.2989/20702620.2024.2390863

Wylie R, Yang CC, Tsuji K (2020) Invader at the gate: The status of red imported fire ant in Australia and Asia. Ecological Research 35(1): 6–16. https://doi.org/10.1111/1440-1703.12076

Yang B, Zhang Z, Yang CQ, Wang Y, Orr MC, Wang H, Zhang AB (2022) Identification of species by combining molecular and morphological data using convolutional neural networks. Systematic Biology 71(3): 690–705. https://doi.org/10.1093/sysbio/syab076

Zengeya TA, Faulkner KT, Mtileni MP, Wilson JRU (2025a) Lessons and challenges in creating alien species lists: Insights from South Africa's national reports on the status and management of biological invasions. NeoBiota 101: 203–222. https://doi.org/10.3897/neobiota.101.162932

Zengeya TA, Faulkner KT, Mtileni MM, Winzer LL, Kumschick S, McCulloch-Jones EJ, Miza-Tshangana SA, Robinson TB, Sifuba A, Engelbrecht W, van Wilgen BW (2025b) A checklist of alien taxa for South Africa. bioRxiv, 2025-05. https://doi.org/10.1101/2025.05.22.655507

Supplementary material 1

Notes containing additional methodological details and R scripts

Authors: Laura Fernández Winzer, Katelyn T. Faulkner

Data type: docx

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.104.162310.suppl1

Supplementary material 2

BOLD-Watch list comparison

Authors: Laura Fernández Winzer, Katelyn T. Faulkner

Data type: csv

Explanation note: List comparison output between BOLD and Watch List, species in Excel file are species from the watch list that have a molecular record in South Africa.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.104.162310.suppl2

Supplementary material 3

BOLD-Inspection Data comparison

Authors: Laura Fernández Winzer, Katelyn T. Faulkner

Data type: csv

Explanation note: Species list comparison output between the BOLD and the Inspection list comparison. The output presents a list of species from the inspection list which have a molecular record from BOLD in South Africa.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.104.162310.suppl3

Supplementary material 4

BOLD-BODATSA comparison

Authors: Laura Fernández Winzer, Katelyn T. Faulkner

Data type: xlsx

Explanation note: Output showing a list of plant species that have a molecular record in the BOLD database for South Africa but that are not present in BODATSA.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.104.162310.suppl4